PROBABILISTIC CHARACTERIZATION OF SIMULTANEOUS NERVE IMPULSE SEQUENCES CONTROLLING DIPTERAN FLIGHT

ROBERT WYMAN

From the Departments of Zoology and Biophysics, University of California, Berkeley

ABSTRACT A probabilistic method of analysis of spike trains is presented which provides a complete statistical description of spike sequences and allows the elucidation of some of the properties of the neural interconnections producing the output patterns. The flight motor system of the blowfly, *Calliphora terraenovae,* is analyzed by this method. Individual motor units show large, nonserially correlated, cycle-to-cycle variations in frequency superimposed upon long term frequency trends. These trends are apparently not generated by averaging the cycle-to-cycle variations in input excitation over a long time period. The different motor units share the same short term input excitation and the excitation causing long term trends. Units in different muscles show no preferred phase or latency relationships; they maintain similar frequencies but their phases drift through all possible values. Frequency control without phase control may be accomplished by shared excitation with a total input frequency many times the output frequency. Units in the same muscle maintain strong phase relationships. Constant phase relationships during variations in frequency may, among other models, be due to reciprocal inhibition or a common linearly rising input. Sensory feedback cannot account for the degree of phase or frequency regulation shown. Thus central patterning of the output sequence is necessary, as in the locust, and the two flight systems can be considered as integradable evolutionary variations.

INTRODUCTION

In the nervous system most information processing occurs at sites in neurons where the electrical signals may have a continuously variable intensity. Because of limitations in current techniques it is not practical, in insects, to record from many of these sites. Yet for many types of experiment insects are the most suitable animals (5), even though the experimenter is often limited to the recording of sequences of axon or muscle spikes. The mechanisms regulating the timing of the spikes cannot themselves be observed, but an analysis of their timing can be made to yield information about the regulatory processes. The analysis must necessarily be a probabilistic one since sequences of spike timings will rarely be repeated exactly in experiments

in which the animal is allowed some freedom in its behavior. This paper introduces a probabilistic method of analysis of spike trains.

Information carried by a nerve spike is usually considered to be all-or-nothing in nature, the exact shape and size of the pulse being irrelevant. One way to analyze spike train data is to arbitrarily construct a waveform by defining an amplitude at each point in time. For instance, the waveform may have a value of zero between spikes and the value of a delta function at the times of occurrence of spikes (3). This construction is useful because the waveforms thus generated have properties which allow the application of many of the standard schemata of mathematical analysis (usually generalizations of Fourier analysis). Another approach, the one described below, is to accept that the only information needed to describe a spike sequence is a list of the times of spike occurrence, separate lists being kept for each axon.

A Probabilistic Representation. Consider two neurons, A and B, from which simultaneous records are made. In any one recorded sequence the times of occurrence of A spikes are $A_1, A_2, A_3, \cdots, A_n$ and of B spikes $B_1, B_2, B_3, \cdots, B_m$. A complete probabilistic characterization of a large number of these sequences is the joint probability density function:

$$p(A_1, A_2, A_3, \cdots, A_n, B_1, B_2, B_3, \cdots, B_m) \qquad [1]$$

which gives the probability that any particular sequence of spike timings will occur. In this paper it is assumed that the timing of any spike is related most strongly to the timing of those spikes that have occurred in a short time period before it. In fact only spans of 11 interspike intervals are considered, although effects lasting more than ten intervals in the experiments will be mentioned. The function [1] can then be restricted to include only spikes in this span:

$$p(A_i, A_{i-1}, A_{i-2}, \cdots, A_{i-11}, B_j, B_{j-1}, B_{j-2}, \cdots, B_{j-11}) \qquad [2]$$

where B_j is the B spike just preceding A_i. The time of occurrence of A_i is set equal to zero.

It is common statistical procedure to estimate a density like [2] from many repetitions of an experiment. An alternative technique would be to use only one long record but consider each different time period encompassing 11 spike intervals as a repetition. A stationary process is roughly one in which the statistical properties of the process are the same at all times; an exact definition may be found in Parzen (9). Probability theory proves that for a stationary process statistics taken from one long record will be the same as those computed from many repetitions of an experiment. The data recorded from insect flight are clearly not stationary. The pattern generated at the beginning and end of a flight is different from that generated in the middle of a flight (see Fig. 2*b* in reference 17). By analyzing only records from the mid-portions of a flight, this type of non-stationarity may be

avoided. The average frequency of spikes vary in different experiments; the correlation coefficient, used in this study, is a statistic which is normalized for frequency. Long-term trends in the frequency of impulses occur during a flight; the effects of this can be determined (Fig. 4) or the data can be broken up into pieces without trends and each piece analyzed separately (Fig. 6). Taking the precautions mentioned, densities will always be estimated from long records rather than from many experiments.

Three elementary manipulations of the lists of spike times give the following useful quantities. The first is the *interval* between spikes in the same unit:

$$a_i = A_{i+1} - A_i. \qquad [3a]$$

The second is the interval between spikes on different neuronal lines:

$$l_i = A_i - B_j \qquad [3b]$$

where B_j is the B spike immediately preceding A_i; $B_j = \sup \{B_k | B_k < A_i\}$.[1] The interval l_i is termed the *latency* of spike A_i with respect to B spikes. The third is a particular ratio of the two already defined:

$$\varphi_i = \frac{l_i}{b_j} = \frac{A_i - B_j}{B_{j+1} - B_j} \qquad [3c]$$

where B_j is the B spike immediately preceding A_i. This quantity φ_i is termed the *phase* of spike A_i with respect to the B spikes. In the results section of this paper approximations to marginal probability densities derived from [2] by the changes of variables defined in equations [3*a*], [3*b*], and [3*c*], and integration will be used to analyze sequences of spikes recorded from the flight muscles of Diptera.

The Flight Motor System of Diptera. In most quickly acting neuromuscular systems the arrival of a nerve impulse at a muscle fiber initiates a contraction of that fiber. The muscles that work the wings of the larger and more primitive insects are of this type and are termed neurogenic muscles. In contrast, some of the muscles of smaller and more highly evolved insects are set into an active state by the receipt of a nerve impulse, but the muscle actually contracts only after it is also stretched (8). This type of muscle is called myogenic. In dipteran flight muscle one nerve impulse may cause an active state which lasts several tenths of a second during which the muscle unit may be stretched thirty times, a contraction resulting from each stretch. During flight in the Diptera the whole thorax becomes a vibrating compartment. Probably the thorax is initially set into vibration by a single twitch of a neurogenic muscle (2). Thereafter the myogenic muscles contract each time they are stretched by the vibrating thoracic box, if they

[1] The function sup { } defines the least upper bound (supremum) in the sequence defined by the bracketed expression.

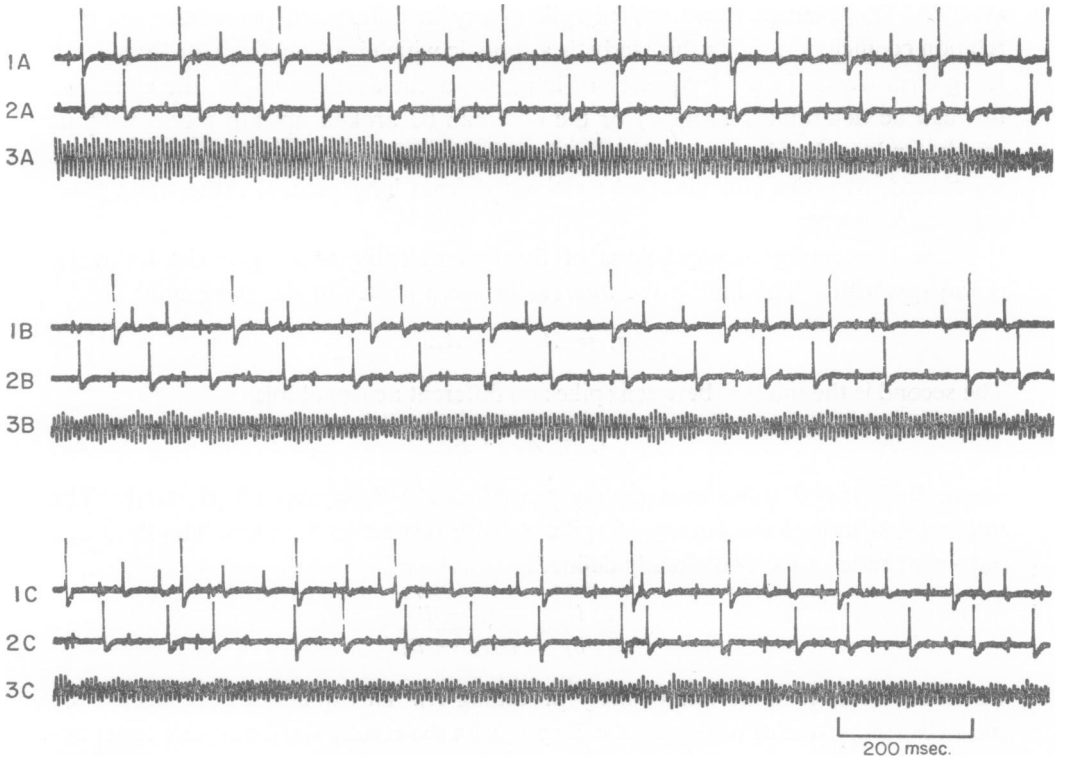


FIGURE 1 Oscilloscopic trace of muscle spikes and wingbeat record. Line 1A (continued through lines 1B and 1C): Three different units in the right dorsal longitudinal muscle. Line 2A, 2B, and 2C: Two units of left DLM (large spikes). Line 3A, 3B, 3C: Record of thoracic movement. Notice the shifting phase relations between units to different muscles. In line 1A the spikes occur in the order large, medium and small in every cycle, but their exact timing with respect to each other varies considerably. The 2 spikes of line 2A, 2B, and 2C occur in strict alternation, one spike is slightly, but with absolute consistency, smaller than the other. The small unit in line 1C has a somewhat lower frequency than the other 2 units. The first 4 cycles of 1C show the sequence L,M,S with the small unit occurring very late in each cycle; it missed the fifth cycle altogether occurring then very early in the 6th cycle, after which the sequence L,S,M persists. In line 1B a very rare event has occurred, a simultaneity of the small and medium spikes in the 3rd cycle at the point of changeover in sequence from L,S,M to L,M,S. Note that wingbeat frequency is about 20 times that of any muscle spike.

have been set into an active state by receipt of a nerve impulse within the last quarter second or so. In this way the timing of the muscle contractions, and thereby the pattern of motion of the wings, is determined more directly by the mechanical properties of the vibrating thorax than by the pattern of nerve impulses to the muscles.

METHODS

Each blowfly, *Calliphora terraenovae,* was attached dorsally to a support which concontained a motion transducer used to record vibrations accompanying the wingbeat. Fine insulated wires were inserted into the thoracic muscles and the output monitored with an oscilloscope. The animal was mounted in front of a wind tunnel and the muscle spikes, which occur one-to-one with the motor neuron spikes, were recorded during tethered flight. Further details of the recording technique may be found in a previous paper (17). The electrical signals were recorded on an Ampex SP300 tape recorder. Interesting portions of the record were played back and filmed with a Grass kymograph camera (Fig. 1). The location of spikes on the film was then measured with a (Gerber Scientific Instrument Co., Hartford, Connecticut) GDDRS-38 digital data reduction system. This digitizer has an accuracy of 0.01 inch; a film speed was used so that 0.01 inch = 1 msec. The sequence of times of occurrence punched on IBM cards by the plotter was used as the input to a program for an IBM 7090. Interval *vs.* time of occurrence plots (Fig. 3) were made by a (California Computer Products, Inc., Anaheim, California), plotter attached to an IBM 1401.

RESULTS—FREQUENCY ANALYSIS

The Pattern of Impulses to a Single Motor Unit. Records containing over 12,000 wingbeat cycles (at frequencies between 120 and 160 CPS) were considered sufficiently long for useful analysis. The wingbeat rate varied from 18 to 28 times the frequency of nerve impulses, giving from 450 to 750 nerve impulses to each muscle unit per analyzed flight. Fig. 1 shows a typical record taken from two electrodes and a mechanoreceptor. Five different motor units are distinguishable; the maxima of the wingbeat record furnish another sequence of timings. In this experiment two of the three units in the right dorsal longitudinal muscle were found to have received exactly the same number of nerve impulses, while the third unit received 6 per cent fewer. Five to six per cent differences in numbers of impulses are also found in comparisons between right and left dorsal longitudinal muscles. The dorsoventral muscles may have up to 14 per cent fewer impulses than the dorsal longitudinal muscles.

Under the change of variable to intervals (equation [3*a*]), the function [2] becomes:

$$p(a_i, a_{i-1}, \cdots, a_{i-10}, b_j, b_{j-1}, \cdots, b_{j-10}).$$

The marginal density which gives the probability of finding different values for the a_i^{th} interval is:

$$p(a_i) = \int_0^\infty \int_0^\infty \cdots \int_0^\infty p(a_i, a_{i-1}, \cdots, a_{i-10}, b_j, b_{j-1}, \cdots, b_{j-10}) \cdot da_{i-1}, \cdots, db_{j-10}.$$

Assuming stationarity $p(a_i) = p(a_j) = p(a)$ for all i and j. The density $p(a)$ can be approximated by a histogram of the interspike intervals (Fig. 2). Spike interval

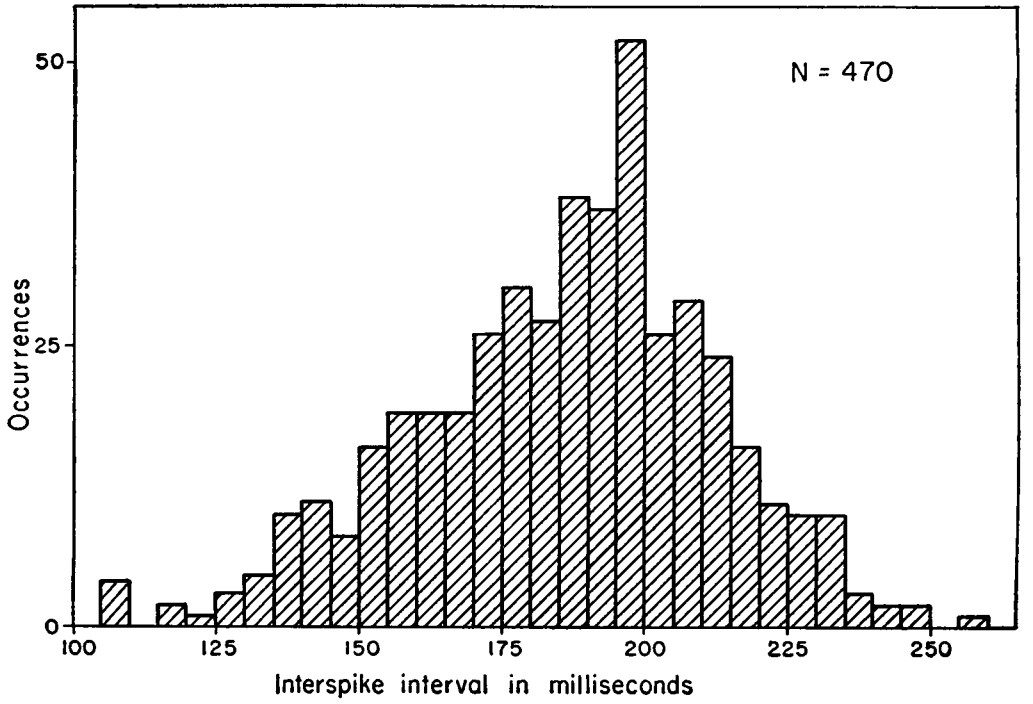


FIGURE 2 Histogram of interspike intervals of a motor unit in the dorsal longitudinal muscle. The mean interval is 185 msec., the standard deviation 26 msec. Although the distribution is roughly bell shaped, it is clearly skewed to the left.

histograms from *Calliphora* units are all roughly bell shaped with standard deviations 14 to 22 per cent of the means. As these histograms do not have the exponential form characteristic of intervals from a random series (*cf.* Fig. 1*a* in reference 18), there must be some mechanism holding the discharge intervals close to a central value. The histograms have a single mode in distinction to histograms from locust flight muscle which are often multimodal (Fig. 9 in reference 14). Thus the dipteran flight motor system does not show the multiple firing characteristic of some insects with neurogenic flight muscles [locusts: (14); Lepidoptera: (6)].

A graphic view of a single spike train may be presented by a time of occurrence *vs.* interspike interval plot. In this type of diagram, A_i, the time of occurrence of spike i, is plotted against $a_{i-1} = A_i - A_{i-1}$, the interval that just preceded spike i. Fig. 3, consisting of two of these plots, shows that the frequency of the units is unsteady; one spike interval may be twice as long as the next. Longer pieces of the plot show that trends lasting many tens of spikes develop so that average frequency may vary by 30 per cent for points separated by 100 nerve spikes.

From Fig. 3 it is not clear whether there is a tendency for adjacent intervals to be similar or dissimilar in length. A statistical study of the order in which intervals

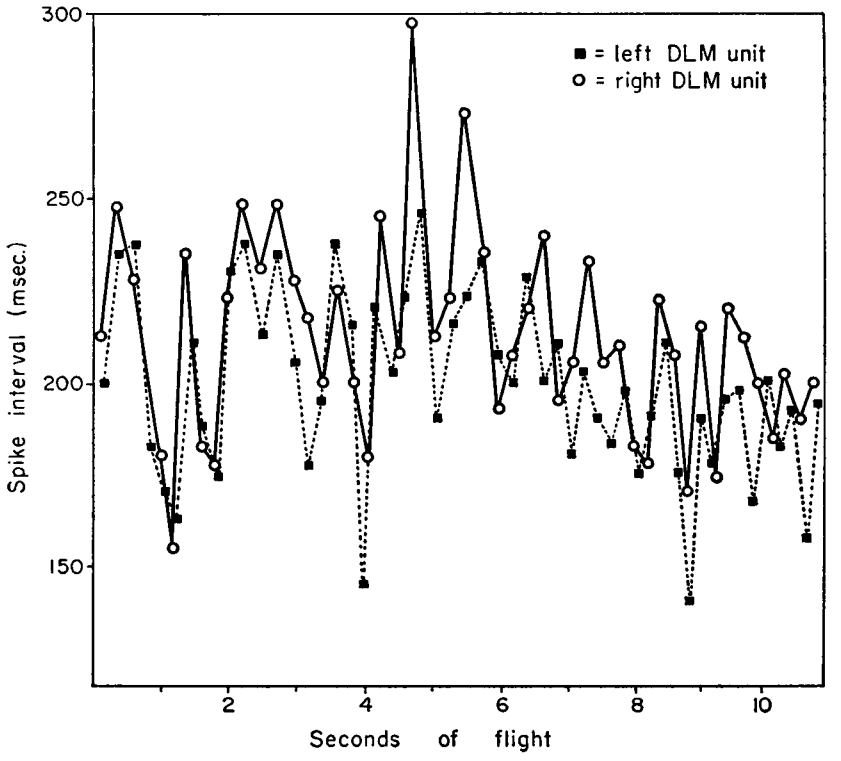


FIGURE 3 Interspike interval *versus* time of occurrence for 2 units. The abscissa of each point is the time of occurrence of a spike and the ordinate is the interval from the previous spike of the same unit. Notice how the units follow each other very closely through changes in frequency.

of different lengths occur provides precise answers for questions of this nature. Consider the densities:

$$p(a_i, a_{i-L}) = \int_0^\infty \int_0^\infty \cdots \int_0^\infty p(a_i \cdots a_{i-10}, b_i \cdots b_{i-10}) \cdot da_{i-1} \cdots da_{i-L+1}, da_{i-L-1} \cdots da_{i-10}, db_i \cdots db_{i-10}.$$

Each of these densities is approximated by the scattergram of each interval *vs.* the L^{th} preceding interval [sometimes called joint interval histograms (11)]. The correlation coefficients for these scattergrams are called the L^{th} order serial correlation coefficient R_L.

$$R_L = \frac{1}{N - L} \sum_{i=1}^{N-L} \left[\frac{a_i - \mu_a}{\sigma_a}\right]\left[\frac{a_{i+L} - \mu_L}{\sigma_L}\right]$$

where

$$\mu_s = \frac{1}{N-L}\sum_{i=1}^{N-L} a_i \qquad \sigma_s = \left[\frac{1}{N-L}\sum_{i=1}^{N-L}(a_i - \mu_s)^2\right]^{1/2}$$

and

$$\mu_L = \frac{1}{N-L}\sum_{i=L+1}^{N} a_i \qquad \sigma_L = \left[\frac{1}{N-L}\sum_{i=L+1}^{N}(a_i - \mu_L)^2\right]^{1/2}.$$

These coefficients have the range -1 to $+1$. When the first order coefficient is negative it indicates that long intervals tend to be followed by short and *vice versa.* With a positive coefficient, long intervals tend to be followed by long ones and short by short. The second order coefficient gives the same information comparing each interval with the interval that follows one after the next. Similarly the L^{th} order coefficient compares each interval with the L^{th} interval following it. When the coefficients stay positive for several orders this means that the data have runs of low frequency activity interspersed with runs of high frequency. The length of the runs is measured by the number of coefficients in a row which are positive. If the 1st 10 orders of serial correlation coefficients are positive but the 11th negative, then the runs tend to be composed of 10 intervals which all stay above, or below, the mean interval. If the first order coefficient is negative and the following ones alternate in sign, then the data also tend to alternate, short interval followed by a long, followed by a short, *etc.* Again the number of coefficients which maintain the alternation in sign gives a measure of the length of these runs of alternation. If the coefficients are all small (insignificant as discussed in Appendix I) then there is no regulation of consecutive interval lengths and the variance is due to random jitter.

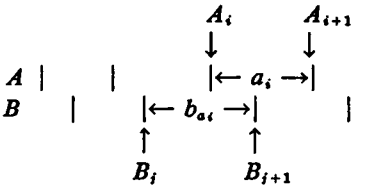
The serial correlation coefficient gives a quantitative measure of how much of the variance in the sequence is generated by serial effects. The square of the correlation coefficient is equal to the fraction of the variance of the interspike intervals which is due to the relation between the length of any interval and the length of a preceding interval. The relation between the intervals may be stated by a regression equation (see Appendix I). Thus although a correlation cofficient of 0.2 may be highly significant (unlikely to have been generated by a random series) it accounts for only 4 per cent of the variance.

Fig. 4*a*, showing the first ten orders of serial correlation coefficient, is typical of any of the flight muscles. The coefficients are all significantly positive, but insignificantly different from each other. That all the coefficients are positive indicates that the flight is composed of runs that are at least 10 intervals long during which the frequency tends to be above or below the mean value. Aside from these long runs there is no tighter regulation on a cycle-to-cycle basis. That the first coefficient is not smaller than the others shows that long intervals are not compensated by being followed by short ones. That the first coefficient is not larger indicates that adjacent intervals are not kept closer in size than are intervals 10 cycles apart. The unsteadi-

ness of the spike train implies that the input to the final spike producing site has large and non-serially correlated cycle-to-cycle variations that remain unsmoothed. If the final spike producing site does not have the capacity for integration with a long time constant (*i.e.*, smoothing) then the extra excitation leading to long term trends must be added in from somewhere else. Possibly this excitation is contained in, or controlled by, the command from higher centers.

Correlated Frequency Variation Between Units. From the interval *vs.* time plot for two contralateral units (Fig. 3) it seems that the interspike intervals of the two units parallel each other in a detailed way through rapid fluctations. Longer pieces of the plot show that the units parallel each other during long terms frequency trends also. Some statistics will now be introduced which will allow these statements to be made more precisely.

We have been considering two units A and B firing at times $A_1, A_2, A_3, \ldots$ and $B_1, B_2, B_3, \ldots$. Define intervals $a_i = A_{i+1} - A_i$ and $b_{a_i} = B_{j+1} - B_j$ where B_j is the largest B still smaller than A_i. We then have two interval sequences a_i and b_{a_i} where each b_{a_i} is the earliest B interval having a portion contiguous with the A interval a_i.



Note that b_{a_i} is not necessarily the i^{th} interval on the line of B spikes although a_i is the i^{th} interval for A spikes. If the B neuron is firing faster than the A neuron as in the illustration below then not all the intervals of the B neuron are included in the b_{a_i} list.

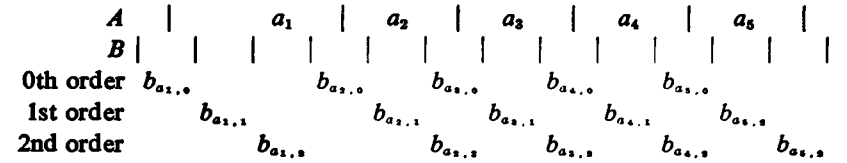
$$\begin{array}{ll} A & \mid a_1 \mid a_2 \mid a_3 \mid \\ B & \mid b_{a_1} \mid \mid \mid b_{a_2} \mid \mid \mid b_{a_3} \mid \mid \mid \end{array}$$

If B is firing more slowly than A then some of the B intervals are on the list more than once.

$$\begin{array}{ll} A & \mid a_1 \mid a_2 \mid a_3 \mid a_4 \mid a_5 \mid a_6 \mid \\ B & \mid b_{a_1} \mid b_{a_2} \,\&\, b_{a_3} \,\&\, b_{a_4} \mid b_{a_5} \,\&\, b_{a_6} \mid \end{array}$$

A scattergram of a_i *versus* b_{a_i} can be drawn and regression lines calculated. The correlation coefficient quantifying the amount of variance explained by these regressions ranges between -1 and $+1$. A positive value indicates that the two units have their periods of high and low frequency together, a negative coefficient indicates that a high frequency period in one unit occurs when the other unit has a low frequency.

It is possible for a unit to follow another in frequency but only after a delay of several intervals. To find the effect after a delay of L intervals the intervals a_i must be correlated with the L^{th} B interval following b_{a_i}, denotable by $b_{a_i,L}$. Let B_j be the largest B still less than A_i then $b_{a_i,L} = B_{j+L+1} - B_{j+L}$. The b_{a_i} previously defined is $b_{a1.0}$ in this notation.



Correlations can then be computed between the a_i list and the $b_{a_i,L}$ lists. The lagged correlation coefficient or L^{th} order coefficient is defined as follows:

$$r(L) = \frac{1}{N-L} \sum_{i=1}^{N-L} \frac{(a_i - \mu_a)(b_{a_i,L} - \mu_b)}{\sigma_a \sigma_b}$$

where

$$\mu_a = \frac{1}{N-L} \sum_{i=1}^{N-L} a_i \qquad \mu_b = \frac{1}{N-L} \sum_{i=1}^{N-L} b_{a_i,L}$$

and

$$\sigma_a = \left[\frac{1}{N-L} \sum_{i=1}^{N-L} (a_i - \mu_a)^2 \right]^{1/2} \qquad \sigma_b = \left[\frac{1}{N-L} \sum_{i=1}^{N-L} (b_{a_i,L} - \mu_b)^2 \right]^{1/2}$$

The zeroth order coefficient is the unlagged coefficient described for a_i *vs.* b_{a_i}. It

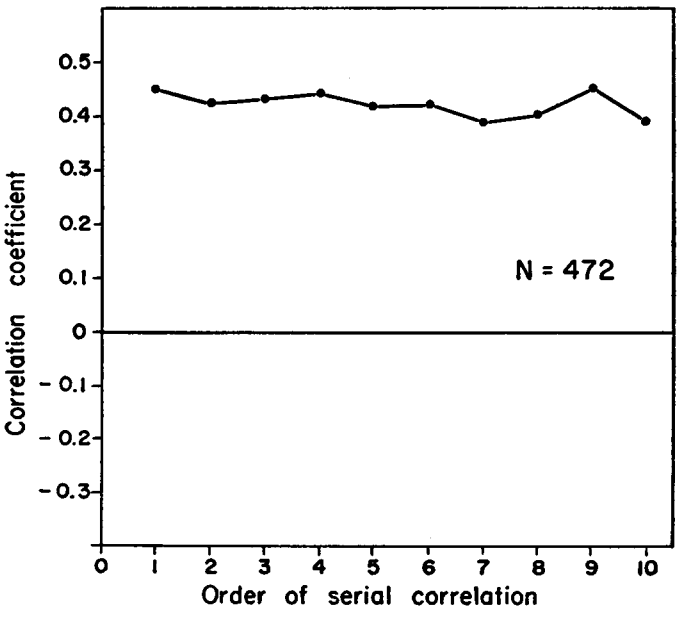


FIGURE 4*a*

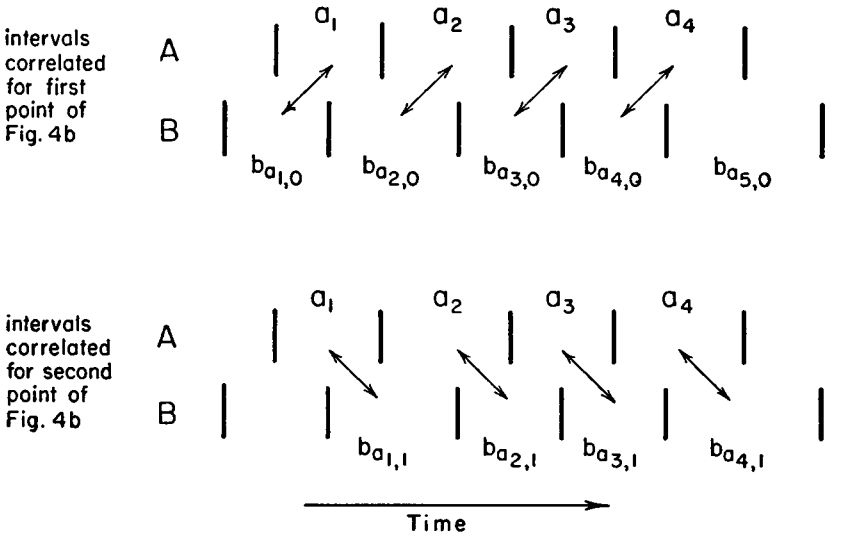


FIGURE 4*b*-1

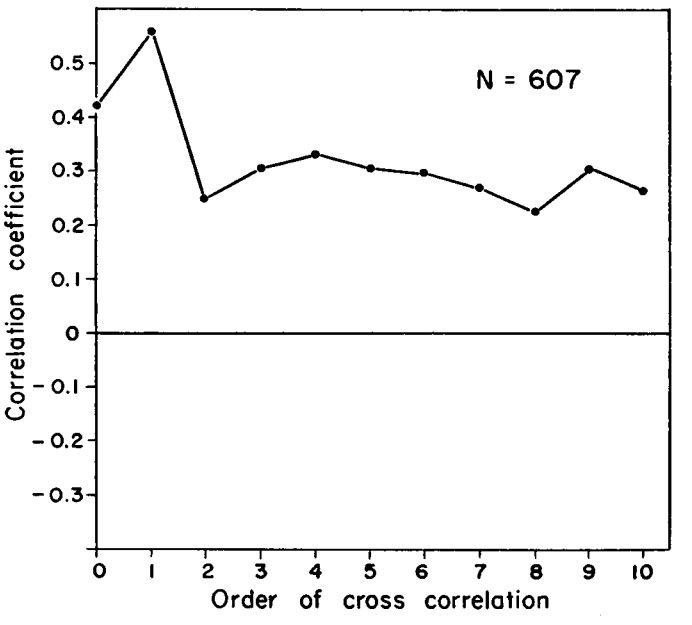


FIGURE 4*b*-2

FIGURE 4 **(*a*) Serial correlogram of interspike intervals to a muscle unit. The first 10 orders of serial correlation coefficient are all high indicating frequency trends lasting longer than 10 spike intervals. (*b*) Frequency correlogram comparing two units. The first two orders are the highest, confirming the cycle-to-cycle frequency following seen in Fig. 2. The highness of the coefficient out to at least 10 orders shows that the long term frequency trends of the two units also follow each other.**

often happens that the zeroth order coefficient is not the largest. In that case the frequency dependence of the two units has its maximum effect delayed by the number of intervals that equals the order of the highest coefficient.

One unit may follow another in frequency not with a delay of a set number of intervals, but after a constant time delay. This constant delay may encompass, at different times, different numbers of intervals depending on the frequency changes of the unit. In this case the frequency following after a delay would not show up as a single high coefficient, but the effect would be spread out, causing several coefficients to be high. In cases of this nature clearer results are obtained by using a form of correlation that is a function of time rather than the present one which is a function of interval number. Such a form for both auto- and cross-correlation is presented in Appendix II.

A typical frequency correlogram comparing 2 units in shown in Fig. 4*b*. Its significant features are the same as is found when comparing any two myogenic units, whether in the same or different muscles. All the coefficients here are positive but the first two coefficients are about twice as large as the others. These first two high correlations confirm the detailed frequency following, on a cycle-to-cycle basis, of the two units. In the discussion of sincle cells the cycle-to-cycle frequency irregularities were attributed to an input which had large and non-serially correlated variations. Since the cycle-to-cycle variations in output are similar for each cell the further implication can be made that the output cells share the same cycle-to-cycle changes in input excitation, and that the effect of this excitation is about the same for each of the cells.

The frequency correlogram gives some further information regarding the utilization of this shared input. The first correlation point in Fig. 4*b* is derived by comparing intervals of *A* and *B* which have a contiguous portion, but where the *B* interval starts before the *A* interval does. The second correlation point compares intervals where in each comparison the *A* interval starts first, although the *A* and the *B* intervals still have a contiguous portion. That the second point is higher than the first means that the closest correlation occurs when each *A* interval is compared to a *B* interval which starts later. This implies that when the output frequency of both units changes the *A* unit's frequency changes first. Under the assumption that output frequency variation is caused by changes in input excitation, it follows that the *A* unit either receives, or reacts to, changes in excitation before *B* does. By comparing each unit to every other, the order in which the different units change their frequency can be determined. If the different units pass excitation to each other in chains (or branched chains) then their hierarchy can also be discovered by this method. The *Calliphora* records do not indicate any regular order in which the units change their frequency nor any hierarchy for the passing of excitation.

We have seen that the high first two correlation coefficients indicate that pairs of units go through similar cycle-to-cycle frequency variations. Correspondingly the

high cross-correlation coefficients of orders 3 to 10 indicates that the units also follow each other through long term frequency trends. The implication from this is that the excitation leading to long term trends is shared by the different cells, and that the effect of this excitation is about the same for each of the cells. Thus, at least during the uncontrolled flight in this set of experiments there is no evidence of differential frequency control, in which one unit's excitation is raised while another's is lowered. In the section on the output of single cells it was argued that the excitation causing cycle-to-cycle variation was distinct from that causing long term trends. In this section it has been argued that both types of excitation are shared, with approximately equal effects, by the different cells.

Correlations between spike frequency and wingbeat frequency are also high and positive, but somewhat smaller than spike frequency correlations between units. Thus in an experiment where the spike frequency correlation coefficients were all about 0.7, the spike to wingbeat frequency correlations were all about 0.6. Also, in every correlation between spike and wingbeat frequencies it was found that the wingbeat changed frequency after the spike frequency did.

RESULTS—PHASE AND LATENCY ANALYSIS

Relations Between Units in Different Muscles. In a previous study (17) it was reported that all phase relationships occurred between simultaneous trains of spikes in different muscles. To complete this statement it should be known whether, all phases being possible, some occur more frequently than others. Fig. 5*a* is a histogram of the phases of spikes in one unit with respect to spikes in a different muscle. If the sequences of spikes in the two units are not related, then any phase should be as likely as any other and the histogram should be flat over its distribution from 0 to 1. In this case (Fig. 5*a*) the probability (X^2 test) is 0.4 that a random sample from a flat population would give a no better fit. Thus there are no marked phase preferences between units in different muscles.

Fig. 5*b* is a typical histogram of the latencies of a unit in one muscle (*B*) with respect to a unit (*A*) in another. This histogram is clearly not flat, there being fewer large latencies than small. This comes about because the latency of a *B* spike can not be longer than the concurrent interspike interval of *A;* at latest the *B* spike can occur at the end of an *A* interval in which case the *A* interspike interval and the *B* latency are equal. Thus while small *B* latencies can occur during either long or short *A* intervals, long *B* latencies can occur only during long *A* intervals.

To predict the shape of the latency histogram under different conditions this dependence of the distribution of *B* latencies on the distribution of *A* interspike intervals must be examined. We shall define a spike train (*B*) which has no preferred latencies with respect to another train (*A*) as one in which (*a*) a *B* spike is as likely to occur during one fraction of an *A* interval as during any other equal

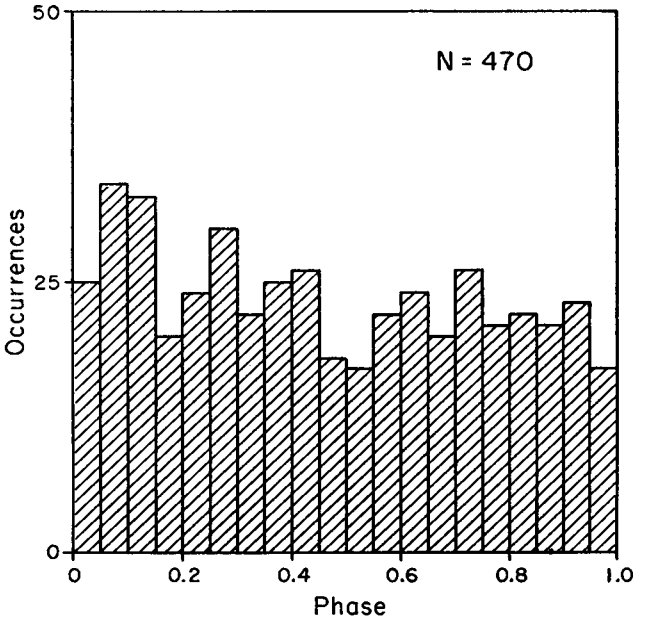


FIGURE 5a

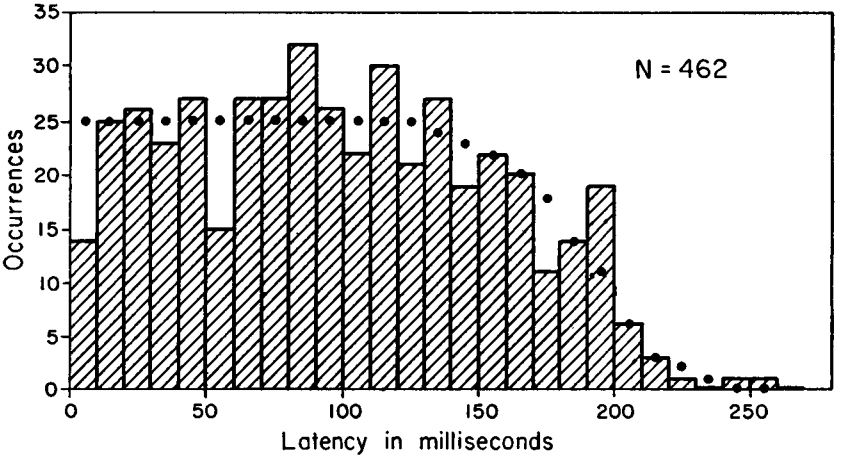


FIGURE 5b

FIGURE 5 (*a*) Histogram of the phases of spikes to a unit in the left DLM as a fraction of the concurrent interval between spikes to a unit in the right DLM. A χ^2 test shows that this distribution does not differ significantly from random. (*b*) Histogram of the latencies of a right DLM unit's spikes after spikes to a left DLM unit (same two units as in Fig. 5*a*). The latency from left to right unit spikes can be short whether the interval between spikes to the left side is long or short. The latency can be long however only when a long interval between spikes to the left side occurs. Thus the histogram should fall off at high values as it does here. If there were no latency regulation between the units the histogram should have the shape of the dotted line.

fraction, *i.e.*, the density function of the phases is flat; and (*b*) the phase of a *B* spike is independent of the length of the concurrent *A* interval.

Let: $f(l)$ be the probability density function of the latencies of the *B* spikes from the *A* spikes that precede each of them,

$g(\varphi)$ be the probability density function of the phases of the *B* spikes in the concurrent *A* intervals,

$h(a)$ be the probability density function of these concurrent interspike intervals of *A*,

$s(a, l)$ be the joint probability density function of *A* interspike intervals and *B* latencies,

$t(a, \varphi)$ be the joint probability density function of *A* interspike intervals and *B* phases.

The two densities $s(a, l)$ and $t(a, \phi)$ are derived from the same events but are expressed in different variables. Hence under the rules for change of variable:

$$s(a, l)\, da\, dl = t(a, \varphi) \frac{\partial(a, \varphi)}{\partial(a, l)}\, da\, dl$$

In the case under consideration the *B* phases are independent of the *A* intervals so that:

$$t(a, \varphi) = h(a)g(\varphi)$$

or

$$s(a, l)\, da\, dl = h(a)g(\varphi) \frac{\partial(a, \varphi)}{\partial(a, l)}\, da\, dl$$

Now the latency of any spike is equal to its phase multiplied by the value of the concurrent interval between *A* spikes:

$$l = \varphi a. \text{ Hence } \frac{\partial(a, \varphi)}{\partial(a, l)} = \frac{1}{a}$$

and

$$s(a, l)\, da\, dl = h(a)g(l/a) \frac{1}{a}\, da\, dl$$

a has its range from 0 to infinity, so:

$$\left[\int_0^\infty s(a, l)\, da\right] dl = \left[\int_0^\infty h(a)g\left(\frac{l}{a}\right) \frac{1}{a}\, da\right] dl$$

Also in the case without preferred latencies the density function of the phases is flat:

$$g(\varphi) = 1 \qquad 0 \leq \varphi \leq 1 \qquad \text{or} \qquad g\left(\frac{l}{a}\right) = \begin{matrix} 0 & a < l \\ 1 & a \geq l \end{matrix}$$

so

$$\left[\int_0^\infty s(a, l)\, da\right] dl = \left[\int_l^\infty \frac{1}{a} h(a)\, da\right] dl$$

but the marginal density

$$\int_0^\infty s(a, l)\, da = f(l)$$

so

$$f(l)\, dl = \left[\int_l^\infty \frac{1}{a} h(a)\, da\right] dl$$

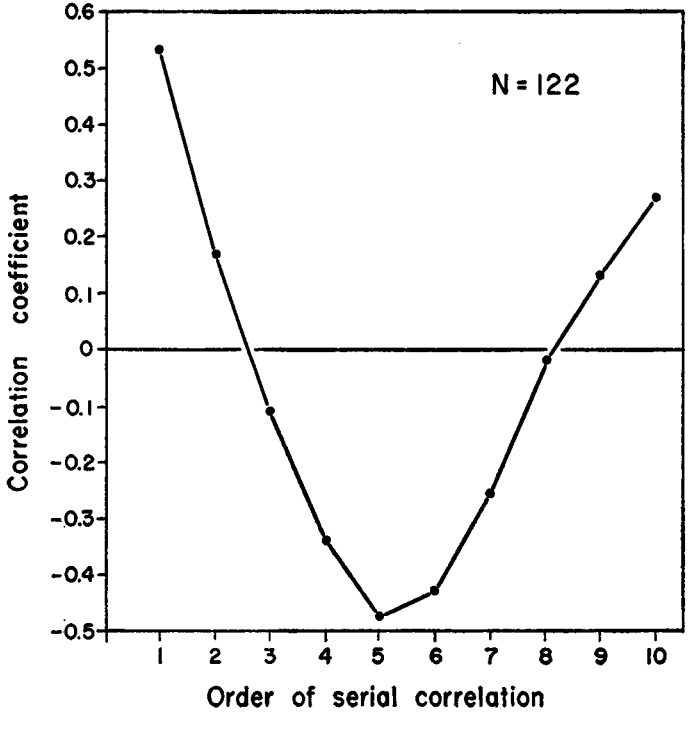
or

$$f(l) = \int_l^\infty \frac{1}{a} h(a)\, da.$$

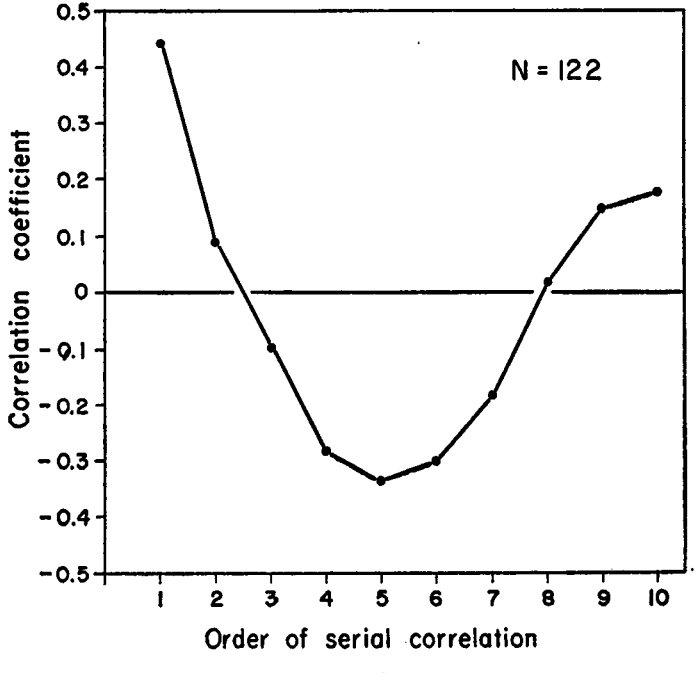
The dots in Fig. 5*b* represent the expected density of latencies for a train of 462 spikes occurring without preferred latencies with respect to a train of spikes having the interval histogram of Fig. 2. The actual spike train is not significantly different from the theoretical (X^2 test, $P = 0.3$). Thus in these comparisons between units in different muscles the timing of the spikes do not show preferred phases or preferred latencies.

Two units in different muscles maintain almost the same frequency, but appear to be otherwise independent since they have no phase or latency dependences. If the units are independent then there should be a pattern in the sequence by which phases (or latencies) follow each other. In any short piece of record the phases will be similar (because the two spike trains are maintaining almost equal frequencies) but there will be a slow drift of phase relationship (because the two spike trains are not maintaining exactly the same frequency). The slower unit will gradually come later and later in the intervals between faster unit spikes until it is near the end of an interval and it then skips an interval and appears next with an early phase. The data from *Calliphora* have so much jitter that this process is not obvious to visual inspection. However, if several orders of serial correlation coefficient are computed on the phases or latencies of a process like this the coefficients themselves should start large and positive for the first order, gradually become small, and then negative, perhaps going through several cycles of positivity and negativity before the jitter damps out the oscillation. Figs. 6*a* and 6*b* are examples of serial correlation on the phases and latencies from one of the *Calliphora* preparations. These two units have nearly the same frequency but the correlograms show that each drifts completely across a cycle of the other approximately every eleven cycles.

In a previous section it was argued that the input which caused the large cycle to cycle variation in interspike interval length was shared by the different motor units. In this section it was shown that in motor units to different muscles this shared input must not cause any phase or latency preferences. One cell could be driving all the motor units by producing an excitatory postsynaptic potential (epsp) in each motor neuron that was sufficient to set off a spike in each of the cells. In this case either (*a*) all the motor units should receive their excitation at the same



6a



6b

FIGURE 6*a* and *b* (*a*) Serial correlogram of phases. (*b*) Serial correlogram of latencies. The oscillation of the correlation coefficient here is caused by oscillation in the sequence of phases (or latencies) correlated. The period of oscillation of the phases (or latencies) is about equal to the period of oscillation of the correlogram.

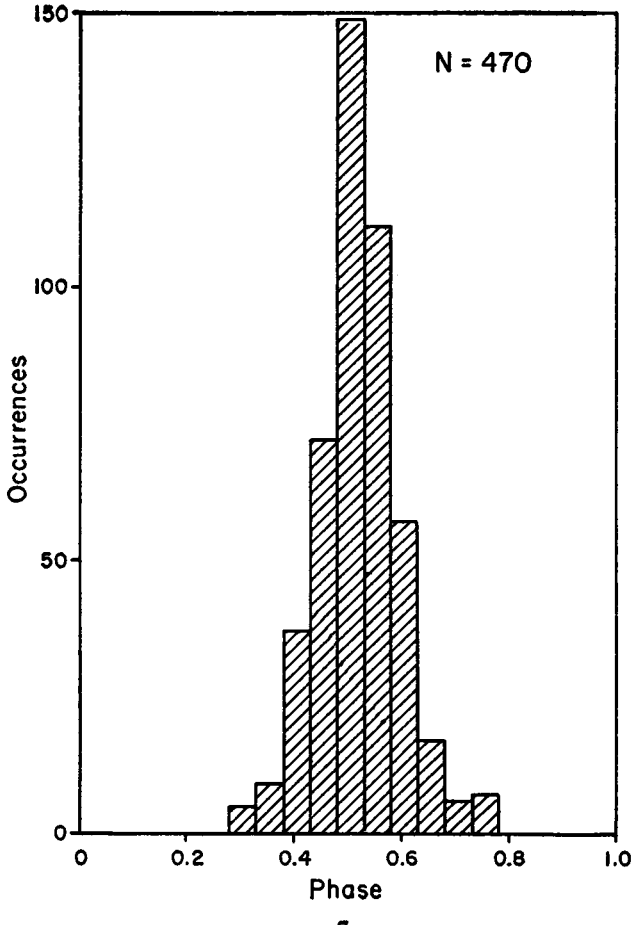
time and thus fire simultaneously; or (*b*) there might be different conduction and synaptic delays from the driver to the different followers resulting in different latencies from the driver to each of the followers. This latter, however, would become evident as the followers would show preferred latencies when compared with each other. A modification of this model that would explain the frequency coupling but that would not cause a phase, or latency, dependence is the following. The driver fires at some rate many times that of the output of any of the follower cells, but causing a smaller epsp for each firing, and the threshold of the follower cells is unequal. Then by summation of the epsp's one cell might fire once for every 10 input pulses, while another with a slightly higher threshold might fire only every eleven input pulses. If the frequency of the driver varied widely this would not cause a latency or phase dependence and would cause the phenomenon of slow drift of phase and latency seen in Figs. 6*a* and *b*.

An equivalent model to the last would have not one driver firing at a high frequency, but many driving cells, all being excitatory with small epsp's to each of the motor cells. If each of these driving cells had a firing time and frequency that was irregular and independent of the others then the summed effect on the motor cells should be large, short term, non-serially correlated fluctuations such as those seen in the highly unsteady firing of the units in Fig. 3.

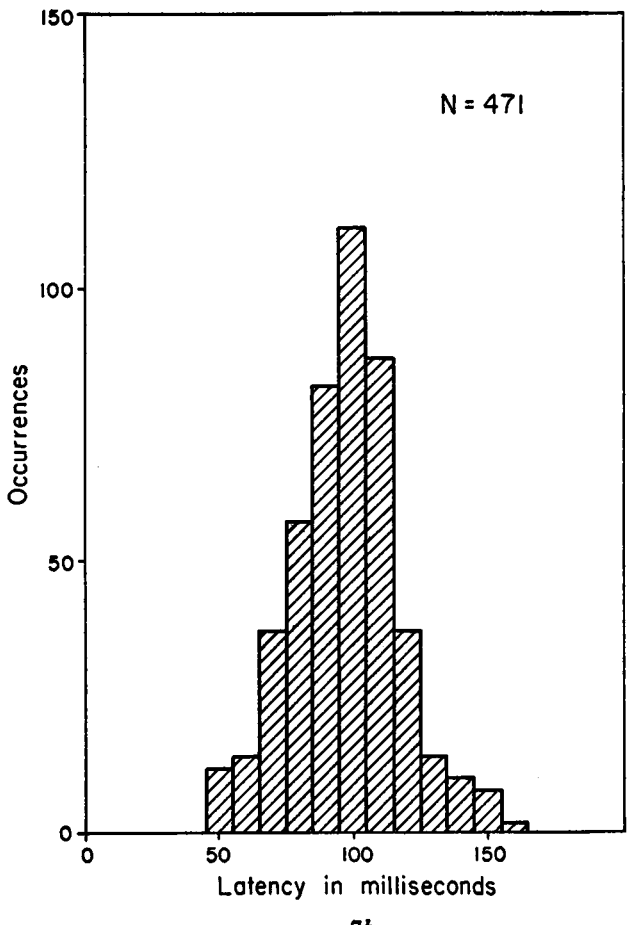
Relations between units in the same muscle. A contrasting situation is found when recordings are made from two units in the same muscle. Here there is a frequency following of about the same strength as between units in different muscles, but there are also strong phase and latency preferences. Fig. 7*a* is a phase histogram comparing units to the same muscle and Fig. 7*b* is the latency histogram for the same comparison. Notice the absolute exclusion of simultaneity or near simultaneity. These two units maintained a strict alternation for over 700 spikes, never coming within 1/5 of a spike interval of each other.

Since both preferred phase and latency relationships show up it should be determined whether this is because one unit, say *A*, tends to fire actually with a specified phase with respect to the sequence of the other unit's (*B*) spikes, or because *A* tends to fire with a constant delay after *B* spikes. For example if the mean interval between *B* spikes is 100 msec. a peak may occur at 0.5 on the phase histogram either because *A* spikes tend to come midway between two *B* spikes, whatever the *B* interval is, or because *A* spikes tend to occur 50 msec. after a *B* spike, whatever the *B* interval is; the latter would cause the phases to vary with the concurrent *B* interval, but the phases would still tend to bunch up around 0.5.

A method of distinguishing between these two possibilities is the following. For two units *A* and *B* the latency of *A* spikes with respect to *B* spikes is, as before, $l_i = A_i - B_j$ where B_j is the largest *B* still less than A_i. The concurrent *B* interval is then $b_j = B_{j+1} - B_j$. The phase of the *A* spike is l_i/b_j. Two correlations are calculated, (*a*) *A* latency *versus* concurrent *B* interval and (*b*) *A* phase *versus* concurrent



7a



7b

FIGURE 7 (*a*) Histogram of phases of one unit in the left dorsal longitudinal muscle compared with another in the same muscle. These two units never fired more closely to each other than 1/5 of a cycle. (*b*) Histogram of latencies of one unit in the left DLM compared with another in the same muscle (the same units as in Fig. 7*a*). Contrast with figures 5*a* and *b* which treat units not both in the same muscle.

B interval. If spikes of unit *A* tend to fall in constant phase with respect to intervals of unit *B* then the phase should not depend on the length of the concurrent interval between *B* spikes, thus giving a zero correlation of *A* phase with *B* interval. However to maintain a constant phase, the latency of *A* must increase with the length of *B* intervals, giving a correlation of latency and interval. On the other hand if there is a latency preference, then *A* latency should have a zero correlation with *B* intervals; but *A* phase must be smaller for long *B* intervals to maintain the same latency; hence *A* phase will be negatively correlated with *B* interval.

For different motor units in the same muscle the above test indicates very clearly that the fundamental control is a phase regulation. Thus for the units of Fig. 7 phase is not significantly correlated with concurrent interval (correlation coefficient = 0.065); while the latency is strongly and positively correlated with concurrent interval (correlation coefficient = 0.719). It would be easy to construct a model if a latency preference were the fundamental phenomenon. Conduction, synaptic, and other delays could be interposed between the driver and follower neurons; differences in the length of the delay to the various follower neurons would appear as latencies between the follower neurons. To construct a model in which the units tend to maintain constant phase, independent of frequency, is a more difficult problem. One possible solution is to have the phased neurons be mutually inhibitory to each other. For further information on this model see references (10) and (4). Another method by which two (or more) cells may be made to maintain a constant phase relationship is the following. Let two cells receive the same excitatory input which causes the membrane potential to rise linearly to a set value. Give the two different thresholds. It can be seen from Fig. 8 that the frequency of firing is determined by the rate of oscillation of the excitation, but that the phases will tend to remain approximately constant independent of frequency.

The following line of argument suggests a possible significance for this regulation of phase. The power delivered by each muscle contraction depends on the frequency of nervous impulses to the muscle, and especially the time since the last impulse. After impulses stop coming the power delivered gradually declines (7). If all the units in a muscle received nerve impulses at the same time the power delivered by each contraction might be significantly different at the beginning of an interspike interval than at the end. By activating the different units in the same muscle at different times in the cycle, one unit may at some time be relatively weak, but another unit will be freshly reexcited. Thus the power delivered from all the units taken together can be smoothed over the 20 to 30 contractions that occur before a nerve impulse is repeated.

DISCUSSION

Comparison of Neurogenic and Myogenic Flight Systems. The best known neurogenic flight system is that of locusts. In locusts different units in the

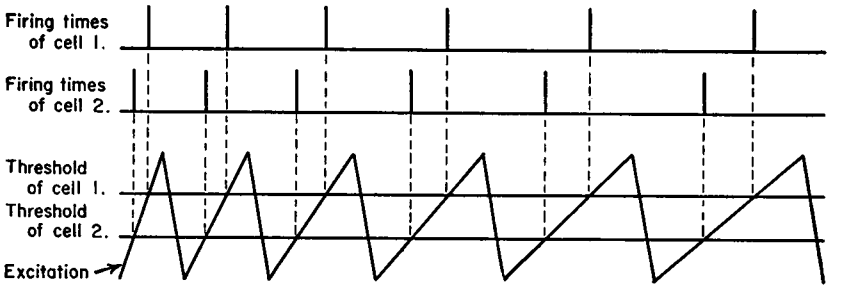


FIGURE 8 If two cells have the same linearly rising input, but different thresholds, their phases will tend to remain constant independent of frequency.

same muscle receive impulses simultaneously, different muscles are activated at specific phases relative to each other, the motions are directly timed by the nervous impulses (16), and the sensory feedback is exactly timed by the motions (15). The locust would appear to require a mechanism of pattern generation very different from a dipteran where there are no phase relationships between impulses to different muscles, between nerve impulses and motions, and therefore between motor impulses and sensory feedback. The locust could, in a reflex loop, use the timing of a sensory input, say from the wing stretch receptors, to time its motor output. That the locust does not depend on this mechanism has been shown (13, 18). The normal output pattern is still generated when the proprioceptive input is random or varying in phase (18), or completely absent (13).

The blowfly system also regulates phase and frequency of different units within a muscle. Since the motor neuron output does not time the motions, proprioceptive feedback cannot be an agent in phase control. Since the frequency correlation between nervous units is larger than the correlation between nervous units and wingbeat, and because the frequencies of the motor units follow each other with a lesser time lag than does the wingbeat, it can be concluded that proprioceptive feedback from the wing motions is not sufficient to cause the degree of frequency correlation observed between units. Thus, as in the locust, intraganglionic interconnections are required for phase and frequency regulation.

The sensory input in locusts is not completely unused. Input from the stretch receptors, for instance, acts to excite the intraganglionic system so that the frequency of its output is increased (Fig. 9 in reference 18). The striking result is that this input need not affect the phase of the output units (Fig. 10 in reference 18). In *Calliphora* the analogous phenomenon is that units in different muscles share a common input but the input is of such a nature that it does not regulate phase.

The major remaining difference between the two systems is that the locust controls the phase of all its units while *Calliphora* only controls the phase of units within the same muscle. Perhaps there is an essential difference between phase controlling interconnections and non-phase controlling interconnections. However,

the intramuscle control in *Calliphora* is really an intermediate case. Here the phase regulation restricts the occurrence of impulses to a fraction of the cycle, but this fraction is neither small, as in the locust, nor the whole cycle as in fly inter-muscle comparisons. This intermediate case makes it reasonable to suggest that the two mecanisms may grade smoothly into each other.

APPENDIX I

The distribution of the serial correlation coefficient is discussed by Anderson (1) for the case when the observations are distributed normally. Table I gives the 1 per cent and 5 per cent confidence levels for R for several N's. When the number of intervals is greater than 30, a normally distributed t may be calculated as

$$t = \frac{(N-1)R_L + 1}{\sqrt{N-2}}$$

Confidence levels on t may then be found in any table of the normal distribution. This is a correct approximation for the first order coefficient and also for the L^{th} order if N and L have no common divisors. When L and N do have common divsors, slight errors are introduced if N is not large.

When the spike interval histogram is not Gaussian somewhat different confidence levels are required. The following test due to Wald and Wolfowitz (12) is valid for a time series with any cumulative distribution function. They compute the statistic

$$C_L = \sum_{i=1}^{N-L} (a_i - \mu)(a_{i+L} - \mu) \quad \text{where} \quad \mu = \frac{1}{N}\sum_{i=1}^{N} a_i$$

As the sample N becomes large this statistic becomes normally distributed for any L that is not an integral divisor of N. The mean of C_L is $E(C_L) = -S_2/N-1$ and the variance is

$$\sigma^2(C_L) = \frac{S_2^2 - S_4}{N-1} + \frac{S_2^2 - 2S_4}{(N-1)(N-2)} - \frac{S_2^2}{(N-1)^2}$$

where

$$S_2 = \sum_{i=1}^{N} (a_i - \mu)^2 \quad \text{and} \quad S_4 = \sum_{i=1}^{N} (a_i - \mu)^4$$

The significance of the normalized variable

$$t_L = \frac{C_L - E(C_L)}{\sigma(C_L)}$$

can be tested with any table of the normal distribution.

The serial correlation coefficient can be used as a predictor. Knowing the i^{th} interval, a_i, the simplest regression equation to predict the $i + L^{th}$ interval a_{i+L}, can be written:

$$(a_{i+L} - \mu_L) = k(a_i - \mu_a). \qquad [1]$$

The best (least mean square error) estimate of k in this equation is $k = R_L(\sigma_a/\sigma_L)$. σ_a is the standard deviation of terms 1 to N-L of the series and σ_L is the standard deviation of terms $L + 1$ to N of the same series. Thus unless the first L terms of the series are

TABLE I

SIGNIFICANCE LEVELS FOR THE SERIAL CORRELATION COEFFICIENT

	Positive correlation		Negative correlation	
Significance level	5 per cent	1 per cent	5 per cent	1 per cent
N				
75	0.173	0.250	−0.199	−0.276
300	0.092	0.131	−0.098	−0.137
500	0.073	0.103	−0.077	−0.107
1000	0.051	0.073	−0.054	−0.075

grossly different from the last L terms, σ_L will closely equal σ_s, or $\sigma_s/\sigma_L = 1$ and $k = R_L$. In that case R_L can be used as the best linear predictor in equation [1].

The correlation coefficients also give the best linear predictors in more compound regression equations. For instance in a linear equation predicting any interval from the two preceding ones:

$$a_i = k_1 a_{i-1} + k_2 a_{i-2}$$

the best fitting k_1 and k_2 are:

$$k_1 = \frac{R_1 - R_1 R_2}{1 - R_1^2}; \quad \text{and} \quad k_2 = \frac{R_2 - R_1^2}{1 - R_1^2}$$

where R_1 and R_2 are the 1st and 2nd order serial correlation coefficients.

In practice only the correlation coefficients themselves need be computed. Regression equations can then be chosen to include those variables having a significant correlation and the regression coefficients can then be calculated from the correlation coefficients. Estimates of coefficients in more complicated regression equations should only be attempted when long series of data are available because, as terms are added to the equation, these regression coefficients become very sensitive to sampling error.

APPENDIX II

An approximation to the probability density for the occurrence of an A spike at any given time (τ) before or after some other A spike is:

$$p(\tau) = \frac{1}{2N+1} \sum_{j=-N}^{N} p(A_i - A_{i-j}) \quad \text{where} \quad A_i - A_{i-j} = \tau \quad \text{for all } j.$$

This is the sum of the probabilities that the first, second, third, *etc.*, spikes after or before any A spike has occurred at time τ. This is equivalent to the estimate of the autocorrelation function given by Gerstein and Kiang (3).

$$\varphi_{ff}(\tau) = \sum_{k=1}^{N} \sum_{i=1}^{N} \delta(t_k - t_i - \tau)$$

where each t_k and t_i represents the time of occurrence of a spike.

Similarly, an approximation to the probability density for the occurrence of a B spike at any given time (τ) before or after an A spike occurrence is:

$$p(\tau) = \frac{1}{2N+1} \sum_{k=-N}^{N} p(A_i - B_{i-k}) \quad \text{where} \quad B_j = \sup \{B_l \mid B_l < A_i\}$$

$$\text{and} \quad A_i - B_{i-k} = \tau \quad \text{for all } k.$$

This is the sum of the probabilities that the first, second, third, *etc.*, B spikes after or before any A spike has occurred at time τ. This formula is equivalent to the cross-correlation function (3):

$$\varphi_{fg}(\tau) = \sum_{k=1}^{N} \sum_{l=1}^{N} \delta(t_k - t_l - \tau)$$

where each t_k is the time of occurrence of an A spike and t_l is the time of occurrence of a B spike.

I am especially grateful to Professor D. M. Wilson for his guidance which has been most valuable in all aspects of this work. The work was done while on the tenure of a National Institutes of Health predoctoral fellowship. The discussion of expected latency distribution is due in part to Miss Ingrid Waldron. F. R. Cole kindly identified the insect species. Financial assistance is acknowledged from research grants from NIH (NB-03927) and NSF (GB-2116) to Dr. Wilson, and from the University of California (Berkeley) Computer Center.

Received for publication, April 22, 1964.

REFERENCES

1. ANDERSON, R. L., Distribution of the serial correlation coefficient, 1942, *Ann. Math. Stat.*, **13**, 1.
2. BOETTIGER, E. G., and FURSHPAN, E., The mechanics of flight movements in Diptera, 1952, *Biol. Bull.*, **102**, 200.
3. GERSTEIN, G. L., and KIANG, N. Y.-S., An approach to the quantitative analysis of electrophysiological data from single neurons, 1960, *Biophysic. J.*, **1**, 15.
4. HARMON, L. D., Neuromimes: action of a reciprocally inhibitory pair, 1964, *Science*, **146**, 1323.
5. HOYLE, G., Exploration of neuronal mechanisms underlying behavior in insects, *in* Neural Theory and Modeling, 1964, R. Reiss, editor, Stanford, Stanford University Press.
6. KAMMER, A. E., Nervous control of wing movement in lepidopterans, 1964, *The Physiologist*, **7**, 172.
7. MACHIN, K. E., and PRINGLE, J. W. S., The physiology of insect fibrillar muscle, II. Mechanical properties of a beetle flight muscle, 1959, *Proc. Roy. Soc., Series B*, **151**, 204.
8. MACHIN, K. E., and PRINGLE, J. W. S., The physiology of insect fibrillar muscle, III. The effect of sinusoidal changes of length on a beetle flight muscle, 1960, *Proc. Roy. Soc., Series B*, **152**, 311.
9. PARZEN, E., Stochastic Processes, 1962, San Francisco, Holden-Day Inc.
10. REISS, R. F., A theory and simulation of rhythmic behavior due to reciprocal inhibition in small nerve nets, 1962, *Proc. Joint Computer Conference*, Palo Alto, California, National Press.
11. RODIECK, R. W., KIANG, N. Y.-S., and GERSTEIN, G. L., Some quantitative methods for the study of spontaneous activity of single neurons, 1962, *Biophysic. J.*, **2**, 351.
12. WALD, A., and WOLFOWITZ, J., An exact test for randomness in the nonparametric case based on serial correlation, 1943, *Ann. Math. Stat.*, **14**, 378.

13. WILSON, D. M., The central nervous control of flight in a locust, 1961, *J. Exp. Biol.*, **38**, 471.
14. WILSON, D. M., Relative refractoriness and patterned discharge of locust flight motor neurons, 1964, *J. Exp. Biol.*, **41**, 191.
15. WILSON, D. M., and GETTRUP, E., A stretch reflex controlling wingbeat frequency in grasshoppers, 1963, *J. Exp. Biol.*, **40**, 171.
16. WILSON, D. M., and WEIS-FOGH, T., Patterned activity of coordinated motor units, studied in flying locusts, 1962, *J. Exp. Biol.*, **39**, 643.
17. WILSON, D. M., and WYMAN, R. J., Phasically unpatterned nervous control of dipteran flight, 1963, *J. Insect Physiol.*, **9**, 859.
18. WILSON, D. M., and WYMAN, R. J., Motor output patterns during random and rhythmic stimulation of locust thoracic ganglia, 1965, *Biophysic. J.*, **5**, 121.